On the Performance of the DQRUMA with Different Scheduling Algorithms for Multimedia Traffic in an Indoor Environment for W-ATM Networks

Ch.V.Verikoukis, J.J.Olmos
Technical University of Catalunia (U.P.C.)
Signal Theory and Communications Department
Campus Nord-D4, Jordi Girona 1-3, 08034-Barcelona, Spain
e-mail: [cveri, olmos]@xaloc.upc.es

Abstract

The performance of the DQRUMA MAC protocol for multimedia traffic in a realistic indoor environment using Guarantee Bandwidth Minimising Delay (GBMD) scheduling policy is evaluated by simulations in this paper. The efficiency of the GBMD is compared to the corresponding with the FCFS with priorities and the RR with priorities.

The analysis was carried out for a typical multicarrier modulation scheme (OFDM) with QPSK modulation on each carrier. The Hidden Markov Model (HMM) is used for modelling the physical layer of the system, while a type-II Hybrid-ARQ protocol based on punctured R-S codes is applied in the LLC layer.

The performance of each ATM traffic service in terms of mean delay, CLR and the standard deviation the 99.9 percentile and p.d.f. of the delay is presented.

1. Introduction

Wireless ATM (W-ATM) is a promising technology for the future broadband services due to its capacity of providing different communication services, in various application environments, while guaranteeing a previously agreed quality of service (QoS) [1].

A typical W-ATM network follows a protocol layer harmonised with the standard ATM. The wireless air interface consists of the Wireless Physical layer (W-PHY) and Data Link Control (DLC) layer. This last contains a Medium Access Control (MAC) and a Logical Link Control (LLC) sublayers.

The performance of the DQRUMA MAC protocol [2] for multimedia traffic, using the GBMD [3] scheduling policy, in a realistic environment is analysed in this paper. The efficiency of the GBMD within the DQRUMA is compared to the corresponding efficiency with the FCFS with priorities and the RR with priorities. The protocol has been simulated in conjunction with a type-II Hybrid ARQ/FEC protocol based on punctured R-S codes [4] for the W-LLC and an OFDM system for the physical layer in an indoor environment. The study is focused in the uplink since, due to the limited power of the mobile transmitter, it is the worst conditioned of the two links.

Due to the long coherence time of the picocellular mobile channel, the detailed physical layer simulation requires programs that execute too slowly. To overcome this problem, the physical layer is modelled by a Hidden Markov Model (HMM). This method reproduces the statistical behaviour of the physical layer, in terms of error distribution and it describes with a good approach the error distribution of the real sequences [5].

The remainder of the paper is structured as follows. In section 2 a brief description of the physical layer is presented. In section 3 the DQRUMA protocol are reported while in section 4 the three scheduling algorithms are analysed. Sections 5 describes the traffic models that are used in the simulations. In section 6 the simulation model is described while in the last section simulations results are done.

2. Physical Layer

A typical OFDM based physical layer is used in the simulations. It provides a 20 Mb/s wireless link in a bandwidth of 25 MHz operating at 5.2 GHz.

16 sub-carriers are used in parallel with QPSK modulation on each one. An Inverse Fast Fourrier Transformation (IFFT) and an oversampling factor of 4 is used to generate the time domain samples transmitted during one OFDM symbol. An oversampling is necessary to avoid aliasing in the generated signal spectrum. To add the cyclic prefix, some samples at the end of each OFDM symbol are replicated to the beginning of the symbol. The

use of a time domain raised cosine (roll-off=0.5) windowing that lasts 15.6% at the beginning and at the end of each OFDM symbol is necessary to reduce the adjacent channel interference. Thus, the total OFDM symbol (1.665μs) is made up of the windowing time, a FFT period of 1.28 μs and a 160 ns cyclic prefix time. Before being transmitted, the signal passes through a non-linear power amplifier with 3dB back-off.

The modulated signal passes through a picocellular radio channel that has a maximum propagation delay of 100ns modelled by three time-variant propagation paths. To estimate the transfer function of the channel, a 10 QPSK symbol pilot sequence known a priori by the receiver is transmitted with ideal back-off once every 1000 OFDM symbols.

In the receiving end, the signal is demodulated coherently and the bits are detected one by one using minimum distance criterion. Ideal time and frequency synchronism is assumed.

3. D.Q.R.U.M.A.

The DQRUMA is an efficient demand-assign type channel access protocol which is designed for fixed-length packets (ATM cells) considering a time slotted system with no frame reference.

It can be divided into two phases, a Request Access (RA) phase and a packet transmission phase. The up-link (UL) stream is divided in a series of minislots used for requesting access (RA-channel), each one followed by a slot for packet transmission (Xmt channel). When a packet arrives at an empty buffer, the mobile sends, probably in contention with other mobiles, a request access (Xmt-Req) to the Base Station (BS) via the up-link RA-channel. When the BS successfully receives a Xmt-Req from a mobile, it sets the corresponding entry in the Request Table (RT) to indicate that a mobile has a packet to transmit. The BS informs the reception of the Xmt-Req by broadcasting the Access ID over the down-link. When not Acknowledgement (NACK) is received, the mobile is dropped in the backlog mode and retries to transmit after a back-off interval.

Dynamic Access channel slotted Aloha is used as random access protocol for the RA-Channel. Pseudo-bayesian algorithm is used by the mobiles in computing their retransmission probabilities after collisions of the their transmit requests.

According to a desired scheduling policy the BS chooses one of the mobiles that has a non empty Xmt-Req field in the RT and broadcasts the Access ID over the down-link channel. With cell transmission is also includes Piggybacking (PGBK) Request, without contention, to indicate it has more packets in its buffer. If all the Xmt-Request fields in the RT are empty, the BS uses the down link Xmt-Perm channel to announce that the next up-link Xmt-channel will be converted into multiple RA-channels.

When channel effects are considered, in the RA-phase only error detection is applied. Some bits in the Xmt_Req packet are only used to detect errors. A NACK message is sent to the mobile if at least an error is detected by the CRC code in the Xmt-Req packet. Then the mobile is dropped in backlog mode. Since the mobiles are situated at different distances from the base station, their signal powers are not the same. Therefore, in case of a collision event, it may still be possible that a packet with the strongest signal strength captures the receiver.

When transmission happens, the type-II Hybrid ARQ/FEC [4] is applied in order to ensure reliable transmission under the worst channel conditions.

4. Scheduling policies

4.1 GBMD

The GBMD service policy combines the desirable for a scheduling police features of isolation and sharing, providing guaranteed bandwidth for each service while attempting to minimise the system delay.

The novel characteristic of this algorithm is that two distinct techniques are used to determine the order of service for cells. The Guaranteed Service Queue (GSQ) is used as a mechanism for delivering guaranteed bandwidth, while the FCFS server optimises delay performance at other instants in time. In [3] the scheduler checks the buffer of each mobile but when the algorithm is used within the DQRUMA the scheduler checks for non-empty entries of the request table.

This service policy operates in two phases to determine the order of service for cells. In the first phase, the cells that must be served in order to ensure that the requested bandwidth is delivered to each mobile are identified. These mobiles place an entry in the GSQ and they are served in FCFS order. It is noted that only mobiles with a non empty entry in the request table can place an entry in the GSQ. Whenever the GSQ is empty in a time slot, the server enters in the second phase. In this phase the oldest cell among all mobiles of the same priority class that has a non-empty entry request table is served. This accomplished by comparing the time of cell generation.

The basic philosophy of this scheduling algorithm is that cells are served through the GSQ only when required. At all other times, cells are served in FCFS order within the same class. This technique relies on a counter that is associated with each mobile; a counter reaching to zero

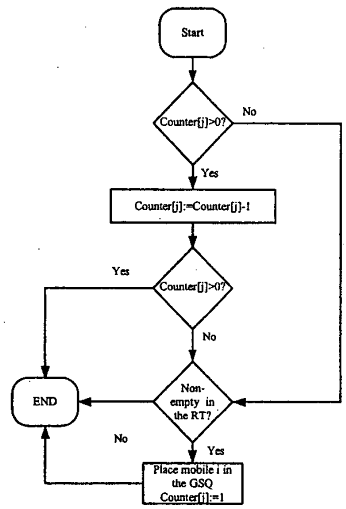

Figure 1 Algorithm performed in each time slot to serve guaranteed cells of mobile *i*.

signifies that a cell from the mobile, if there is an entry in the RT, is to be placed in the GSQ. Initially, counters for all the mobiles are set to *1*. At the beginning of each time slot, for each mobile, if counter j is greater than *0*, then counter j is decremented by a fixed value, r_j. The value r_j is calculated based on guaranteed requested by mobile j and is equal to the fraction of server capacity that is allocated to mobile j. If counter j is less than or equal to *0*, after decrementing the counter, an entry from the user j is to be placed in the GSQ. If the mobile has a non-empty entry to the request table, then an entry is placed in the GSQ and the counter j is reset to *1*. If for this mobile there is no entry to the request table or if counter j is greater than *0*, then the algorithm ends for mobile j in the current time slot. Once the counters for all mobiles are updated, and entries are placed in the GSQ, then a cell from a user that have an entry in the GSQ is served, if any present. If there are no entries in the GSQ, then the scheduler serves the cell with the oldest cell. The counter of this flow is then reset to *1*.

4.2 FCFS with priorities and RR with priorities

For every priority class i the base station maintain a request table called *RT_i*. In each time slot the base station searches for a non-zero entry in the table with the higher priority is executed. Starting from priority class 4 (CBR-voice) and going down to priority class 1 (ABR-WWW) the scheduler checks if all entries in the *RT_i* are zero, then the table for priority class $i-1$ is examined. Generally only if all entries in a priority class i are zero, the table with the lower priority is examined. If at some priority class there exists a non-zero entry then the corresponding mobile terminal will be allowed to transmit a cell in the next time slot. If a non-empty field does not exist in all the request tables the base station informs via the downlink channel that the next slot will be converted in minislots. After that a new search in the next time slot always starts from the higher priority table. Within a priority class FCFS and RR scheduling mechanisms are applied to determine which mobile is allowed to transmit.

When FCFS is utilised the tails of the delay distribution are minimised, while with RR isolation between the flows is attained.

5. Traffic Models

Traffic models for voice, video, data and WWW are used in order to evaluate the performance of the above scheduling algorithms in the DQRUMA. Each of the models correspond to one of the ATM traffic services.

Voice traffic-A voice traffic can be characterised by the two-state Markov process [6]. The users alternates between the ON and OFF states which correspond to the talkspurts and idle periods of speech. In the ON state voice packets are generated at a constant bit rate of 64 kbits/s, while no packets are generated in the OFF state during to the silence period. The activity and the silence intervals have exponential distribution with mean 1s and 1.35s, respectively. A voice cell is dropped when it remains in the buffer longer than 10ms.

Video traffic-The video traffic represents real time VBR traffic. To incorporate large peaks in the rate occurs when scene changes take place, due to the complete update of the image a two-state model is used.

In the first state a source is modelled using a method belonging to the class of discrete-time batch Markov arrival processes (D-BMAP) [7]. In this case each source is modelled as superposition of 10 identical independent ON-OFF sources, each of them generating traffic at constant rate during the ON period. The time interval in each period have geometrical distribution. Mean rate of 256 kbits/s and standard deviation of 128kbits/s is assumed. In the second state cells are generated with probability equal to the peak rate of the source. A traffic source remains in this state for a geometrically distributed time with mean a video frame duration. Maximum delay allowed for a cell generated by a video source is 10ms.

Data traffic-The traditional Poisson process cannot capture the fractal-like behaviour of the data traffic [8]. Self-similar models show that the traffic has similar statistical properties at a range of time scales. In [8] long range dependent traffic with Hurst parameter $H=(3-\alpha)/2$ is obtained by the superposition of a number of ON-OFF sources in which the ON-OFF periods have a Pareto type distribution with infinite variance. In the simulation the α-

value in the Pareto distribution is equal to 1.2, while mean data rate of 256 kbits/s and standard deviation of 128kbits/s are used. A data cell is discarded if not transmitted after 500 ms.

WWW traffic-A typical WWW browsing *session* [9] consist of a sequence of *packet calls*. During a packet call several *packets* may be generated, which means that the packet call constitutes of a bursty sequence packets. a packet call corresponds to uploading of a WWW document (e.g. e-mail, TCP ACK etc.). Each packet contains several ATM cells. After the document arrives to the BS, the user consumes certain amount of time for preparing another WWW document. The number of packet calls per session, the time between two consecutive packet call requests, the number of packets in a packet call and the time interval between two consecutive packets inside a packet call are modelled as a geometrically distributed random variable. The session arrival process is modelled as a Poisson process. For the packet size, Pareto distribution with cut-off is used. A maximum packet size of 65535 bytes is allowed while the cell lifetime is 1s.

6. Simulation Model

In the simulation model the base station is located at the centre of a circular cell of unit radius and the position of the mobiles is uniformly distributed in the circle and does not change during the simulation.

Before starting the simulation the received signal power at the base station from each mobile normalised to the received signal power at the border of the cell is calculated. Path loss exponent, that has the typically for picocellular indoor environment value of 2 is assumed. The E_b/N_0 depending on the $E_b/N_{0border}$ is then defined and the corresponding HMM transition matrix is assigned to each mobile. It is underlined that $E_b/N_{0border}$ of 12 dB has been calculated.

At the beginning of each time slot, the state of each transition matrix is updated. Using these transition matrices, the error distribution both in Xmt_Req packets and in transmitted cells is produced. Apart from these matrices, a set of transition matrices called "capture matrices" for values of E_b/N_0, in steps of 1 dB, are also updated at the beginning of each slot. In case of collision an equivalent signal power called "capture power", for the user with the higher signal power between the collided users, is calculated. This "capture power" depends on the power of the other collided users. Based on this value the appropriate HMM "capture matrix" is used to produce the error distribution in the Xmt-Req packet.

Considering that the physical layer is OFDM based, each time slot consists of 17 OFDM blocks. The first of them is used as Xmt_Req packet. In the other 16 OFDM blocks the ATM payload, the ATM compressed header, all the other necessary headers, error detection and error correction bits are carried. A multimedia scenario is used in order to evaluate by simulation the previous described scheduling policies within the DQRUMA.

Table 1 WWW model parameters, up-link

Parameter	Value
Arrival session probability	$8*10^{-7}$
Average number of packet calls in a session	5
Average reading time between packet calls	412s
Average number of packet in a packet call	25
Average interarrival time between packets	10.4ms

7. Simulation Results

In our scenario 60% of the users have data and WWW capability simultaneously, using notebook computer or palmtop. 15% of the users have simultaneously data and voice capabilities. 10% of the users have only voice capability and the remaining 10% only video capability. For all the services 64 kbit/s minimum bandwidth are guaranteed with the GBMD.

The performance of the time sensitive applications is evaluated by presenting plots for the mean delay, the standard deviation, the 99.9 percentile and the p.d.f. of the delay as well as for the CLR. The standard deviation and the 99.9 percentile of the delay are statistical indicators for the magnitude of the delay jitter and they correspond to measures of the width and the tail of the delay density function respectively.

For the data and the WWW traffic the mean delay and CLR are analysed. Since these services are not time sensitive any other delay parameter are not plotted.

As throughput the ratio of the total received cells without errors to the total number of simulated slots is defined. In the throughput the time dedicated in each time slot to the request access packet, an OFDM block, is included. The mean delay is measured in time slots while CLR is the ratio of the total lost cells to the total transmitted cells is defined.

Figure 2 and Figure 3 present the mean delay and the CLR versus throughput for the voice and the video traffic. The efficiency of the RR degrades under heave load and it is worst among the three policies. With the RR higher mean delay values than the other schemes are observed for all the throughput values while higher CLR it can be seen for throughput values higher than 0.78. The performance

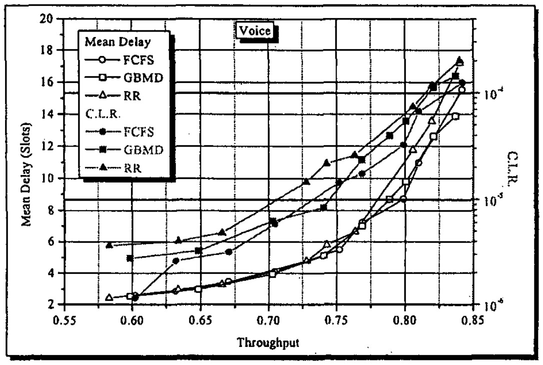

Figure 2 Mean delay and CLR vs. throughput, voice traffic

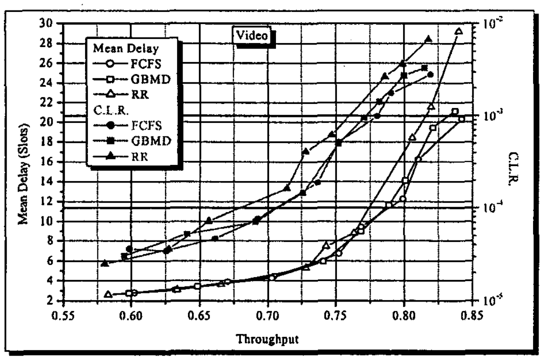

Figure 3 Mean delay and CLR vs. throughput, video traffic

of the GBMD is close to the corresponding with the FCFS.

Figure 4 and Figure 5 compare the standard deviation and the 99.9 percentile of the delay for the voice and the video traffic. For the voice traffic the standard deviation of the delay for a throughput value of 0.84 is 27.33, 23.4 and 22.35 slots for the RR, GBMD and FCFS, respectively. The corresponding delay for the video traffic are 53.03, 41.5 and 41.6 slots. Similarly the 99.9 percentile of the delay for voice traffic are 249 slots for the RR, 224 slots for the GBMD and 220 for the FCFS. It subtracted from these result that the efficiency of the GBMD is better than the RR and is similar to the FCFS.

In Figure 6 and in Figure 7 the p.d.f. of the delay for the voice and the video traffic are depicted. It is evident from these plots that the events with a delay higher than 50 slots appear with lower probability with the FCFS and GBMD in both of the traffic services.

Figure 8 compares the CLR for the data and WWW of the examined scheduling policies. For the data traffic it can be seen that the RR policy performs slightly better than the other policies and the GBMD has the worst efficiency. It is noted that for this type of traffic lost cells

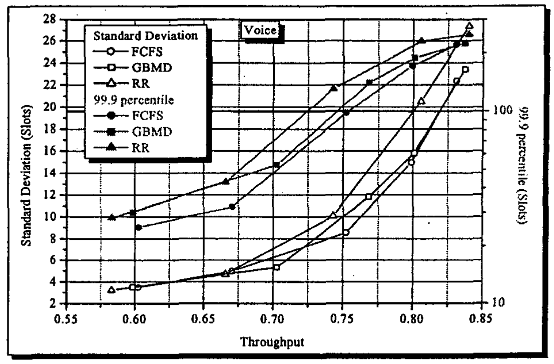

Figure 4 Standard deviation and 99.9 percentile of the delay vs. throughput, voice traffic

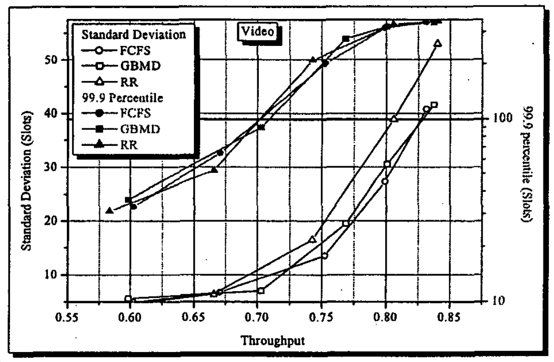

Figure 5 Standard deviation and 99.9 percentile of the delay vs. throughput, video traffic

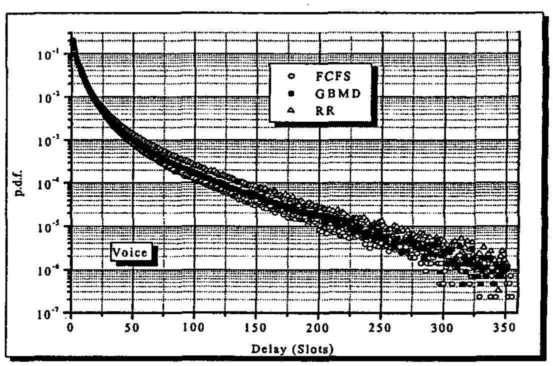

Figure 6 p.d.f. of the voice delay

are not observed for throughput values lower than 0.8. It is obvious from the same plot that lower CLR values of the WWW traffic are achieved when the GBMD is applied. This fact happens due to minimum guaranteed bandwidth is provided by this policy. In Figure 9 the corresponding mean delay for the data and the WWW traffic is compared.

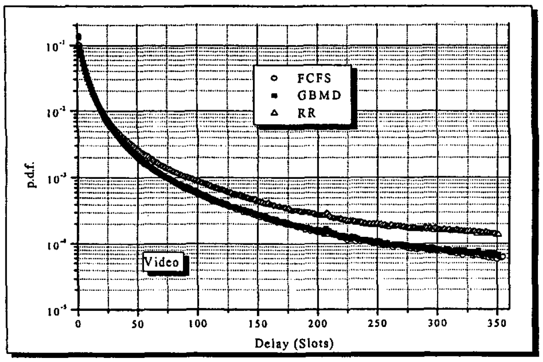

Figure 7 p.d.f. of the video delay

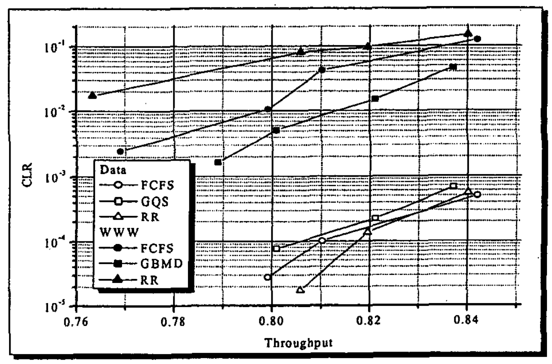

Figure 8 CLR vs. throughput, data and WWW traffic

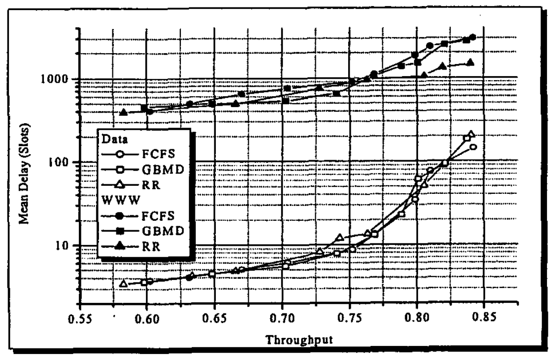

Figure 9 Mean delay vs. throughput, data and WWW traffic

8. Conclusions

The performance of the DQRUMA using GBMD scheduling policy in a realistic indoor environment has been presented in this paper. The efficiency of the GBMD has been compared to the efficiency of the FCFS with priorities and the RR with priorities scheduling policies within the DQRUMA. Simulations results for the mean delay, the delay jitter the CLR as well as for the p.d.f. of the delay of each traffic service have been done.

From the simulations results it can been seen that the performance of the GBMD policy is better than the RR with priorities and it is close to the FCFS with priorities efficiency. Therefore with the GBMD the requested bandwidth of each mobile is guaranteed without any significant degradation in the mean delay, the delay jitter and the CLR of the traffic services.

References

[1] D.Raychaudhuri, "ATM Based Transport Architecture for Multiservices Wireless Personal Communication Network" *IEEE JSAC*, vol. 12, no.8, Oct. 1994.

[2] M.J.Karol, Z.Liuand, K.Y.Eng, "Distributed-Queuing Request Update Multiple Access (DQRUMA) for Wireless Packet (ATM) Networks," *Proc. IEEE INFOCOM'95*, pp. 1224-1231.

[3] P.Pancha, M. Karol "Guaranteeing Bandwidth and Minimising Delay in Packet-Switched (ATM) Networks" *IEEE VTC'1995*, pp.1064-1070, 1995.

[4] Ch.V.Verikoukis, J.J.Olmos, "An Efficient Type-II Hybrid ARQ protocol using Punctured R-S Codes for Wireless ATM Networks," *Proc. VTC'99-Fall*, pp.1725-1729.

[5] Ch.V.Verikoukis, J.J.Olmos "Modelling of a Wireless ATM Physical Layer Based on an OFDM Modulation Scheme Using Hidden Markov Model", *International Workshop on Mobile Communications'99*, Chania, Crete, Greece, 1999.

[6] P.T.Brady, "A model for generating on-off speech patterns in two way conversation," *Bell Syst. Tech. J.*, vol.48, no.7, pp.2445-2447, Sept. 1969.

[7] C. Blondia and O. Casals, "Performance Analysis of statistical multiplexing of VBR sources," *Proc. of INFOCOM 1992*, pp.828-838.

[8] W.E.Leland et all., "On the self similar nature of Ethernet traffic," *IEEE/ACM Trans Networking*, vol.2, pp. 1-15, Feb. 1994.

[9] ETSI TR 101 112 V3.1.0 "Universal Mobile Telecommunications System (UMTS) Selection Procedure for the choice of radio transmission technologies of the UMTS" Nov. 1997.